Astronomical Data Analysis Software and Systems XIII
ASP Conference Series, Vol. 314, 2004
F. Ochsenbein, M. Allen, and D. Egret, eds.

Data Processing at the INTEGRAL Science Data Centre

M. Beck (Mathias.Beck@obs.unige.ch), A. Aubord, P. Bartholdi, V. Beckmann, P. Binko, J. Borkowski, T. Contessi, T. Courvoisier, P. Dubath, K. Ebisawa, M. Gaber, D. Götz, T. Jaffe, D. Jennings, P. Kretschmar, D. Landriu, I. Lecoeur, L. Lerusse, T. Lock, M. Meharga, S. Mereghetti, N. Morisset, N. Mowlavi, S. Paltani, J. Peachey, B. O'Neel, N. Produit, R. Rohlfs, A. Sauvageon, S. Shaw, M. Türler, R. Walter

Integral Science Data Centre, Chemin d'Ecogia 16, 1290 Versoix, Switzerland, http://isdc.unige.ch

Abstract. *INTEGRAL*, the *INTE*rnational *G*amma *R*ay *A*strophysics *L*aboratory of the European Space Agency, was launched on October 17, 2002. *INTEGRAL* operations are performed very successfully and the instruments provide data of good quality. These data are routed to the *INTEGRAL* Science Data Centre (ISDC), where they are processed, archived and finally distributed to the science community. In this paper we present the organisation of the centre and its software development. Some details of the data flow are given along with a description of the core ISDC software libraries.

1. Introduction

The main duties of the ISDC are the reception of the telemetry, the search for Gamma-ray bursts in real-time, the detection of new and transient sources with a delay of a few hours after their observation and the archiving and distribution of the scientific data products.

To fulfill those tasks a team of scientists and software engineers was set up. The software development was performed jointly with the teams who built the instruments on board *INTEGRAL*. The instrument teams provide the so called instrument specific software (ISSW) requiring a deep understanding of the instruments' behaviour. ISDC's focus is the processing infrastructure and the integration of the ISSW into the overall system.

2. The *INTEGRAL* Science Data Centre

The *INTEGRAL* Science Data Centre is attached to the Geneva Observatory and is organised as a consortium of scientific institutes from 12 countries spread over Europe and the U.S. From those institutes the ISDC receives valuable resources in the field of scientific knowledge, data processing experience and software engineering skills.

The ISDC started its activities in 1995 with a total staff of 2 Full Time Equivalents (FTE). At the time of the *INTEGRAL* launch in 2002 the number of staff had grown to some 35 FTE. They were and still are actively supported by numerous members of the instrument teams.

Part of the ISDC system is currently being adapted to be used for the processing of the data from the Low Frequency Instrument on board ESA's Planck satellite. Details can be found in the paper of Türler et al. in this volume.

3. Data Flow at the ISDC

Most of the *INTEGRAL* observation time is spent dithering around the targets of interest. As a result the accumulated data can be broken down into data groups corresponding to distinct periods of pointings with a stable attitude and slews, when the spacecraft changes the orientation. The data organisation and processing are based on these time intervals called science windows.

Some 120 kBit/s of telemetry data are received in real-time at the ISDC. In addition, the data are received a second time after recovery of data lost between the ground stations and the ISDC. For a description of the *INTEGRAL* ground segment please refer to the paper by Walter et al. in this volume.

3.1. Real-time Data Processing

- Detection of gamma-ray bursts
  The real-time data are used to perform a fast search for gamma-ray bursts. A dedicated system has been developed at the ISDC to trigger on these events and to broadcast the information with a minimum delay to a community of subscribed users. Several gamma-ray bursts in the field of view of the IBIS instrument onboard *INTEGRAL* have been detected between the launch and the time of this paper. The first burst was detected in November 25, 2002 and the typical delay between the reception of the telemetry at the ISDC and the reception of the alert by the subscribed clients is a few tens of seconds.
- Search for new or transient sources
  On a time scale of a couple of hours the real-time telemetry is analysed to detect new or transient sources. This task is broken down into several steps. All telemetry data is fed into the pre-processing where nearly all the house keeping and science packets are decoded and stored in the FITS format for further processing. The pre-processing is described in detail in the paper by Morisset et al. in this volume.
  After a few intermediate steps for time format unification, data conversion and automatic calibration, the data is processed by the 'Quick Look Analysis Pipeline' where images are reconstructed. These are then used to search for time varying or new sources.
  The whole processing chain from telemetry reception through the creation of images to the comparison of observed source fluxes and positions with catalogue data is fully automatic. It requires only minimal human monitoring to react to anomalies in the data or the processing software. In case a new or transient source is detected by the ISDC software a manual

inspection of the data is performed. If the detection is confirmed, feedback is provided to the *INTEGRAL* ground segment and to the science community.

3.2. Data Archive

To recover from potential data loss between the ground stations and the ISDC, the telemetry is consolidated by the Mission Operations Centre located in Darmstadt, Germany. There the most complete telemetry data available is written to CD-ROM and sent to the ISDC with a typical delay of two to three weeks after the actual observation.

The consolidated data are used at ISDC to populate the *INTEGRAL* data archive. Besides the telemetry and raw data, high level data products like images, light curves and spectra are generated and stored in the archive. Those products comply to the HEASARC/OGIP standards and may be further analysed using standard tools like XSPEC etc.

In total some 2.5 GB of data are archived per day. This leads to a few TB of data during the planned lifetime of the *INTEGRAL* satellite. All data in the archive are stored on hard disks and thus are permanently available for further use. Once in the archive, the data corresponding to observations from accepted proposals are provided to the corresponding principal investigator. The distribution media available are network access and tape media. The community and the ISDC largely prefer the network access via standard FTP.

After an initial proprietary period of one year all data become publicly available and can be downloaded from the *INTEGRAL* archive via W3browse developed at HEASARC. For a more detailed description of the archive and its data browser please refer to the paper by Meharga et al. in this volume.

4. Organisation of the Software Development

The software development for the ISDC was based on the classical waterfall life cycle as suggested by the ESA PSS-05 software standard. In the early development phase, user requirements, software requirements and architectural design documents were established. Each phase ended with a review of the corresponding documents. The reviews were organised with the participation of ESA as well as the instrument teams, who provided the instrument specific software modules for the ISDC system.

Throughout the development phase and in particular during the time close to the launch, the ISDC participated in many ground tests and as well as instrument calibrations. The data from those activities were extremely valuable to improve the ISDC software and the interfaces between the various parts of the ground segment.

4.1. Configuration Management and Change Control

The full ISDC software system consists of some 380 components. Twenty five of those are software libraries. Please see section 5. for a further discussion of the ISDC Support Software.

The ISDC software is mostly developed on SUN Solaris and, where needed, ported to Linux. Most of the off-site users of the scientific analysis software do have Linux as their preferred operating system.

The individual software components are formally delivered to the ISDC software librarian who registers them. Upon successful completion of compilation and unit tests the software is stored in a central software library. From there modules are checked out for integration, test or operational purposes.

Once a component is sufficiently stable, changes are to be documented in the form of Software Problem Reports or Change Requests. A Change Control Board has been put in place to assess and coordinate the proposed changes.

4.2. Daily Builds

The latest versions of all components are frequently built to detect errors or incompatibilities introduced due to recent changes. In addition the unit test of each component is executed. The output data and log-files of the current test run are compared to a reference output provided by the software developers at delivery time. In this manner, unwanted side effects of changes or bug fixes can be detected early in the cycle.

5. Support Software

Early in the development phase ISDC provided a common set of interface libraries:

- Data Acess Layer (DAL) – see the paper by O'Neel et al. in this volume
- Report Interface Layer (RIL) – unified style of log and error messages
- Parameter Interface Layer (PIL) – following the IRAF parameter style
- A GUI extension to the IRAF parameter files – based on special comment lines in the parameter file

The interfaces provided by the support software are exclusively used by all ISDC components. Care was taken not to allow direct access to low level functions e.g. access to CFITSIO etc.

The support libraries are mostly implemented using the C programming language. As a requirement from some of the instrument teams they also provide FORTRAN90 bindings.

6. Conclusions

The ISDC software system was developed between 1995 and 2002 by a consortium of scientific institutes. Instrument specific software modules were provided by the teams who developed the instruments. The integration of those modules was performed at the ISDC.

Since the beginning of the *INTEGRAL* operations in October 2002, the ISDC system was operated without any major problem. Procedures and tools have been implemented allowing the minor problems to be fixed with a generally short delay. The design of the ISDC system has proven to be robust. In addition it provides enough flexibility to adapt to operational changes.